

PARTIAL TRANSLATION OF JP 11(1999)-18194 A

Publication Date: January 22, 1999

Title of the Invention: MICROPHONE ARRAY APPARATUS

Patent Application Number: 9-170288

Filing Date: June 26, 1997

Inventor: Naoshi MATSUO

Applicant: FUJITSU LTD.

[Claim 2] A microphone array apparatus according to claim 1, comprising a delay unit connected to a front stage of the filter, and a delay calculating part for obtaining a cross-correlation function value between an output signal of the plurality of microphones and the noise source signal, and obtaining a delay amount of the delay unit based on a condition under which the cross-correlation function value becomes maximum.

[Claim 8] A microphone array apparatus having a microphone array in which a plurality of microphones are arranged, comprising:

a signal estimating part for estimating that estimation microphones are arranged in accordance with an arrangement interval of the microphones based on an output signal of the plurality of microphones and a propagation velocity of a sound wave, and outputting an output signal of the estimation microphones together with an output signal of the microphones constituting the microphone array; and

a synchronization adding part for adding the output signal of the microphones constituting the microphone array in the signal estimating part to the output signal of the estimation microphones under the condition that phases thereof are matched with each other.

[Claim 11] A microphone array apparatus according to claim 8 or 9, comprising: a signal estimating part for dividing a sound source direction with respect to the microphone array by a predetermined angle, and

outputting an output signal of the microphones constituting the microphone array and an output signal of the estimation microphones estimated based on the output signal of the microphones constituting the microphone array in accordance with the divided directions; a synchronization adding part for adding the output signals of the signal estimating part under the condition that phases thereof are matched with each other; and a sound source position detecting part for outputting sound source position information based on a maximum value of the output signal of the synchronization adding part.

(Page 5, left column, lines 36–40)

The microphone array apparatus can include a delay unit connected to a front stage of the filter, and a delay calculating part for obtaining a cross-correlation function value between an output signal of the plurality of microphones and the noise source signal, and obtaining a delay amount of the delay unit based on a condition under which the cross-correlation function value becomes maximum.

(Page 5, right column, line 45 – Page 6, left column, line 2)

The microphone array apparatus can include: a signal estimating part for dividing a sound source direction with respect to the microphone array by a predetermined angle, and outputting an output signal of the microphones constituting the microphone array and an output signal of the estimation microphones estimated based on the output signal of the microphones constituting the microphone array in accordance with the divided directions; a synchronization adding part for adding the output signals of the signal estimating part under the condition that phases thereof are matched with each other; and a sound source position detecting part for outputting sound source position information based on a maximum value of the output signal of the synchronization adding part.

(Page 7, right column, line 49 – Page 8, left column, line 6)

Assuming that an output signal from microphones 1-1 to 1-n is $gp(i)$ (where $p = 1, 2, \dots, n$, $i = \text{sample No.}$), a cross-correlation function value $Rp(i)$ with a signal $x(j)$ from a noise source is obtained as presented by the following expression.

$$Rp(i) = \sum_{j=1}^s gp(j+i) * x(j) \dots (13)$$

$\sum_{i=1}^s$ represents addition from $i=1$ to $i=s$, and “s” represents the number of samples for performing convolution operation. The number of samples “s” generally can be set tens of samples to hundreds of samples.

(Page 8, left column, line 15 – left column, line 29)

A value “ip” ($p = 1, 2, \dots, n$) of “i” when an absolute value of a cross-correlation function value $Rp(i)$ obtained by Expression (13) becomes maximum is obtained, and furthermore, a maximum value “imax” of “ip” is obtained. This processing is conducted in accordance with the steps shown in (AI) to (AII) in FIG. 4. More specifically, assuming that imax = initial value (e.g., 0) and $p = 1$ (A1); $Rpmax = \text{initial value (e.g., 0, 0)}$, $ip = \text{initial value (e.g., 0)}$, and $i = 0$ (A2); and a cross-correlation function value $Rp(i)$ by the above-mentioned Expression (13) is obtained (A3).

It is determined whether or not the cross-correlation function value $Rp(i)$ is larger than $Rpmax$. In the case where the cross-correlation function value $Rp(i)$ is larger than $Rpmax$, $Rp(i)$ is set to be $Rmax$ (A5). In the case where cross-correlation function value $Rp(i)$ is smaller than $Rpmax$, $i = i + 1$ (A6). Then, it is determined if $i \leq D$ (A7). When “i” is equal to or smaller than the maximum delay sample number D, the process proceeds to Step (A3). When “i” exceeds the maximum delay sample number D, the process proceeds to Step (A8). In Step (A8), it is determined whether or not “ip” is larger than “imax”. In the case where “ip” is larger than “imax”, “ip” at that time is set to be “imax” (A9). In the case where “ip” is not larger than “imax”, $p = p + 1$ (A10). It is determined whether or not $p \leq n$ (A11). When $p \leq n$, the process proceeds to Step (A2). When this condition is not satisfied,

the search for the cross-correlation function value $R_p(i)$ is completed, and the maximum value “imax” of “ip” in a range of $i \leq D$ is obtained.

Using “ip” and “imax” obtained by the above-mentioned detection of a maximum value, the delay sample number “dp” of the delay unit is obtained by the following expression.

$$dp = imax - ip \dots(14)$$

Because of this, the delay sample number $d1$ to dn of the delay units 8-1 to 8-n are set by the delay calculating part 9.

(Page 13, left column, lines 5-6)

An intended sound can be emphasized by synchronization addition.

(Page 13, left column, lines 30-31)

An output signal of a power at a maximum value can be output as an intended sound emphasizing signal.

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-18194

(43)公開日 平成11年(1999) 1月22日

(51)IntCl. ⁵	識別記号	F I
H 0 4 R 3/00	3 2 0	H 0 4 R 3/00 3 2 0
G 0 1 S 3/802		G 0 1 S 3/802
H 0 3 H 17/00	6 0 1	H 0 3 H 17/00 6 0 1 G
H 0 4 R 1/40	3 2 0	H 0 4 R 1/40 3 2 0 A
5/027		5/027 Z
審査請求 未請求 請求項の数12 O L (全 19 頁)		

(21)出願番号 特願平9-170288

(22)出願日 平成9年(1997) 6月26日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72)発明者 松尾 直司

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74)代理人 弁理士 柏谷 昭司 (外2名)

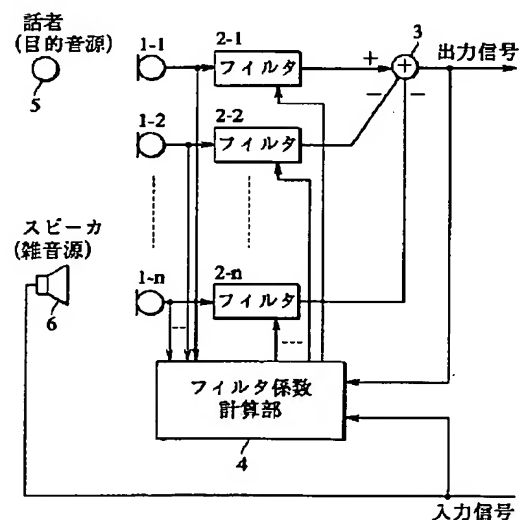
(54)【発明の名称】 マイクロホンアレイ装置

(57)【要約】

【課題】 複数のマイクロホンを配列して構成したマイクロホンアレイ装置に関し、音源位置検出、目的音強調、雑音抑圧等を安定に行う。

【解決手段】 複数のマイクロホン1-1~1-nの出力信号を入力するフィルタ2-1~2-nと、フィルタ2-1を介した基準マイクロホン1-1の出力信号から、フィルタ2-2~2-nを介した他のマイクロホン1-2~1-nの出力信号を減算する加算器3と、複数のマイクロホン1-1~1-nの出力信号と、加算器3の出力信号と、話者5を目的音源とした時のスピーカ6を雑音源とし該スピーカを駆動する為の信号(入力信号)とを基に、フィルタ2-1~2-nのフィルタ係数の更新制御を行うフィルタ係数計算部4とを備えている。

本発明の第1の実施の形態の説明図



【特許請求の範囲】

【請求項 1】 複数のマイクロホンを配列して構成したマイクロホンアレイを有するマイクロホンアレイ装置に於いて、

前記マイクロホンの出力信号を入力するフィルタと、
前記マイクロホンの出力信号と、雑音源信号と、前記フィルタを介したマイクロホンの出力信号の中の基準マイクロホンの出力信号から他のマイクロホンの出力信号を減算した残差信号とを入力し、該残差信号を基にした評価関数に従って前記フィルタの係数を求めるフィルタ係数計算部とを備えたことを特徴とするマイクロホンアレイ装置。

【請求項 2】 前記フィルタの前段に接続した遅延器と、前記複数のマイクロホンの出力信号と前記雑音源信号との相互相関関数値を求めて該相互相関関数値が最大値となる条件を基に前記遅延器の遅延量を求める遅延計算部とを設けたことを特徴とする請求項 1 記載のマイクロホンアレイ装置。

【請求項 3】 前記雑音源信号を、スピーカを駆動する信号としたことを特徴とする請求項 1 又は 2 記載のマイクロホンアレイ装置。

【請求項 4】 前記複数のマイクロホンからなるマイクロホンアレイと共に、雑音源信号を出力する補助マイクロホンを設けたことを特徴とする請求項 1 又は 2 記載のマイクロホンアレイ装置。

【請求項 5】 前記フィルタ係数計算部に於けるフィルタ係数の更新処理に於ける畳み込み演算のフィルタのメモリ値に対して、重み付けを小さくする為の巡回型ローパスフィルタを設けたことを特徴とする請求項 1 乃至 4 の何れか 1 項記載のマイクロホンアレイ装置。

【請求項 6】 複数のマイクロホンを配列して構成したマイクロホンアレイを有するマイクロホンアレイ装置に於いて、

前記マイクロホンの出力信号を入力する線形予測フィルタと、

前記マイクロホンの出力信号を入力して前記線形予測フィルタのフィルタ係数を線形予測分析に従って更新する線形予測分析部と、

前記線形予測フィルタの出力信号の線形予測誤差信号を基に相関係数値を求め、該相関係数値が最大となる値を基に音源位置情報を出力する音源位置検出部とを備えたことを特徴とするマイクロホンアレイ装置。

【請求項 7】 目的音源をスピーカとし、該スピーカを駆動する信号を入力して、前記複数のマイクロホン対応の前記線形予測フィルタに対するフィルタ係数更新の制御を行う線形予測分析部を設けたことを特徴とする請求項 6 記載のマイクロホンアレイ装置。

【請求項 8】 複数のマイクロホンを配列して構成したマイクロホンアレイを有するマイクロホンアレイ装置に於いて、

前記複数のマイクロホンの出力信号と音波の伝搬速度とを基に、前記マイクロホンの配置間隔に従って推定マイクロホンが配置されたと推定し、該推定マイクロホンの出力信号を前記マイクロホンアレイを構成するマイクロホンの出力信号と共に出力する信号推定部と、
該信号推定部の前記マイクロホンアレイを構成するマイクロホン及び前記推定マイクロホンの出力信号の位相を合わせて加算する同期加算部とを備えたことを特徴とするマイクロホンアレイ装置。

【請求項 9】 前記マイクロホンアレイを構成する複数のマイクロホンの配置線上に前記マイクロホンの配置間隔に従って配置した参照マイクロホンを設け、前記信号推定部は、前記マイクロホンアレイを構成する前記マイクロホンの出力信号を基に前記推定マイクロホンの配置位置及び該推定マイクロホンの出力信号を補正する構成を有することを特徴とする請求項 8 記載のマイクロホンアレイ装置。

【請求項 10】 前記参照マイクロホンの出力信号と前記信号推定部で推定した前記参照マイクロホンの配置位置の推定マイクロホンの出力信号との差の誤差信号に対して、聴覚特性に従った重み付けを行って聴覚感度が高い帯域の推定精度を高くする推定用係数決定部を設けたことを特徴とする請求項 9 記載のマイクロホンアレイ装置。

【請求項 11】 前記マイクロホンアレイに対する音源の方向を所定の角度に分割し、該分割した方向対応に、前記マイクロホンアレイを構成するマイクロホンの出力信号及び該出力信号を基に推定した推定マイクロホンの出力信号とを出力する信号推定部と、該信号推定部の出力信号を位相を合わせて加算する同期加算部と、該同期加算部の出力信号の最大値を基に音源位置情報を出力する音源位置検出部とを有することを特徴とする請求項 8 又は 9 記載のマイクロホンアレイ装置。

【請求項 12】 複数のマイクロホンを配列して構成したマイクロホンアレイを有するマイクロホンアレイ装置に於いて、

前記複数のマイクロホンの出力信号を基に音源位置を検出する音源位置検出部と、

前記音源を撮像するカメラと、

該カメラの撮像信号を基に前記音源位置を検出する検出部と、

前記音源位置検出部からの音源の位置情報と、前記検出部からの音源の位置情報とを基に、前記音源の位置を示す音源位置情報を出力する統合判定処理部とを備えたことを特徴とするマイクロホンアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数のマイクロホンを配列して信号処理により音源位置検出、目的音強調、雑音抑制等を行うマイクロホンアレイ装置に関する

る。マイクロホンアレイ装置は、例えば、無指向性の複数のマイクロホンを配列して、目的音強調や雑音抑制等により等価的に指向性を与えることができるものである。又複数のマイクロホンの出力信号の位相関係により音源位置を検出することにより、例えば、テレビ会議システム等に於ける発言者の方向にテレビカメラを自動的に移動して、発言者の音声と共に映像を伝送することができる。又その場合に周囲の雑音を抑制することによって、発言者の音声の明瞭化を図ることができる。又発言者の音声の位相を合わせて加算することにより発言者の音声強調を行うことができる。このようなマイクロホンアレイ装置の動作の安定化を図ることが要望されている。

【0002】

【従来の技術】従来例のマイクロホンアレイ装置として、雑音抑制を目的とした場合、各マイクロホンにフィルタを接続し、雑音成分が最小となるように適応的或いは固定的にフィルタ係数を設定する構成が知られている（例えば、特開平5-111090号公報参照）。又音源位置検出を目的とした場合、各マイクロホンの出力信号の位相関係を求めて、音源方向及び音源までの距離を計測する構成が知られている（例えば、特開昭63-177087号公報又は特開平4-236385号公報参照）。

$$J = e^2$$

として、フィルタ係数 c_1, c_2, \dots, c_r の更新制御を行うものであり、

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \end{bmatrix} = \begin{bmatrix} c_{1old} \\ c_{2old} \\ \vdots \\ c_{r old} \end{bmatrix} + \alpha * (e / f_{norm}) * \begin{bmatrix} f(1) \\ f(2) \\ \vdots \\ f(r) \end{bmatrix} \quad \dots (2)$$

但し、 $0 < \alpha < 0.5$

$$f_{norm} = (f(1)^2 + f(2)^2 + \dots + f(r)^2)^{1/2} \quad \dots (3)$$

に示すものとなる。なお、*は乗算記号、 r はフィルタ次数を示し、又 $f(1), \dots, f(r)$ はフィルタのメモリの値（サンプル単位で遅延させる遅延器の出力信号）を表し、又ノルム f_{norm} は（3）式に示すものとなる。又 α は定数で、フィルタ係数の最適値への収束の速さと精度とを表すものである。

【0006】このようなエコーキャンセラ201に於いては、フィルタの次数が数100となる。そこで、図21に示すマイクロホンアレイを用いたエコーキャンセラが知られている。同図に於いて、211はエコーキャンセラ、212は送受信インタフェース部、214-1～214-nはマイクロホンアレイを構成するマイクロホ

*【0003】又雑音抑制技術としてエコーキャンセラが知られている。例えば、図20に示すように、ネットワーク203に電話機の送受信インタフェース部202が接続され、マイクロホン204とスピーカ205との間にエコーキャンセラ201が接続され、話者の音声をマイクロホン204に入力し、スピーカ205から相手話者の音声を再生することにより、相互に通話を行うことになる。

【0004】その時に、点線矢印経路でスピーカ205からマイクロホン204に回り込む音声、相手電話機に対するエコー（雑音）となる。そこで、減算器206と、エコー成分生成部207と、係数計算部208とを含むエコーキャンセラ201が設けられている。このエコー生成部207は、スピーカ205を駆動する信号を基にエコー成分を生成するフィルタ構成とする場合が一般的であり、減算器206に於いてエコー成分を減算し、その残差を最小とするように、係数計算部208はエコー生成部207のフィルタ係数の更新制御を行うことになる。

【0005】このフィルタ構成のエコー成分生成部207のフィルタ係数 c_1, c_2, \dots, c_r の更新は、既に知られている最急降下法を適用して求めることができる。例えば、減算器206の出力信号 e （エコー成分の残差信号）を基に評価関数 J を、

$$\dots (1)$$

※【数1】

※

ン、215はスピーカ、216は減算器、217-1～217-nはフィルタ、218はフィルタ係数計算部である。

【0007】この場合、マイクロホン214-1～214-nに対してスピーカ215から点線矢印の経路で音声が入力されてエコーとなるから、スピーカ215が雑音源となる。そこで、話者が発音していない場合に、フィルタ217-1～217-nのフィルタ係数 $c_{11}, c_{12}, \dots, c_{1r}, \dots, c_{n1}, c_{n2}, \dots, c_{nr}$ の更新制御は、評価関数を（1）式と同一とすると、

【数2】

$$\begin{matrix} 5 \\ \begin{bmatrix} c11 \\ c12 \\ \vdots \\ c1r \end{bmatrix} = \begin{bmatrix} c11_{old} \\ c12_{old} \\ \vdots \\ c1r_{old} \end{bmatrix} - \alpha * (e / f_{l_{norm}}) * \begin{bmatrix} fl(1) \\ fl(2) \\ \vdots \\ fl(r) \end{bmatrix} \end{matrix} \quad \dots (4)$$

$$\begin{matrix} \begin{bmatrix} cp1 \\ cp2 \\ \vdots \\ cpr \end{bmatrix} = \begin{bmatrix} cp1_{old} \\ cp2_{old} \\ \vdots \\ cpr_{old} \end{bmatrix} + \alpha * (e / f_{p_{norm}}) * \begin{bmatrix} fp(1) \\ fp(2) \\ \vdots \\ fp(r) \end{bmatrix} \end{matrix} \quad \dots (5)$$

但し、 $p=2, 3, \dots, n$

に示すものとなる。

【0008】この場合、(4)式は、複数のマイクロホン 214-1~214-n の中のマイクロホン 214-1 を基準マイクロホンとして、この基準マイクロホンの出力信号を入力するフィルタ 217-1 のフィルタ係数 $c11, c12, \dots, c1r$ について示し、又 (5) 式は、基準マイクロホン以外の他のマイクロホン 214-2~214-n の出力信号をそれぞれ入力するフィルタ 217-2~217-n のフィルタ係数 $c21, c22, \dots, c2r, \dots, cn1, cn2, \dots, cnr$ について示す。そして、減算器 216 に於いては、基準マイクロホンに対応するフィルタ 217-1 の出力信号に対して、他のマイクロホンに対応するフィルタ 217-2~217-n の出力信号を減算する構成を有するものである。

【0009】図 22 は従来例の音源位置検出及び目的音強調処理の説明図であり、221 は目的音強調部、22*

$$r(i) = \sum_{j=1}^n a(i) * b(i+j) \quad \dots (6)$$

で表される。なお、 $\sum_{j=1}^n$ は $i=1$ から $i=n$ まで加算することを示し、又 i は、 $-m \leq i \leq m$ の関係を有

$$m = (\text{サンプリング周波数}) * (\text{マイクロホン間距離}) / (\text{音速}) \quad \dots (7)$$

となる。又 n は畳み込み演算を行うサンプル数で、一般には数 100 となる。

【0012】又 Z^{-da} の遅延器 223 と Z^{-db} の遅延器 224 との遅延サンプル数 da, db は、相関係数値 $r(i)$ の値が最大となる時の i の値より、

$i \geq 0$ の場合、 $da = i, db = 0$

$i < 0$ の場合、 $da = 0, db = -i$

とする。それによって、音源からの目的音の位相が一致されて加算器 226 により加算され、目的音が強調されて出力される。

【0013】

【発明が解決しようとする課題】雑音抑制の為の従来例に於いて、マイクロホンアレイと共にスピーカ等の雑音源を有する場合、目的音源の話者が発声しない場合に、スピーカからの再生音がマイクロホンアレイに回り込ん

* 2 は音源位置検出部、223, 224 は遅延器、225 は遅延サンプル数計算部、226 は加算器、227 は相関係数値計算部、228 は位置検出処理部、229-1, 229-2 はマイクロホンを示す。

【0010】目的音強調部 221 は、 Z^{-da} と Z^{-db} との遅延器 223, 224 と、遅延サンプル数計算部 225 と、加算器 226 とを含む構成であり、又音源位置検出部 222 は、相関係数値計算部 227 と、位置検出処理部 228 とを含む構成である。この音源位置検出部 222 は、相関係数値計算部 227 により、マイクロホン 229-1, 229-2 の出力信号 $a(i), b(i)$ の相関係数値 $r(i)$ を求め、位置検出処理部 228 により、相関係数値 $r(i)$ が最大となる時の i の値 i_{max} により音源位置を求めて、遅延サンプル数計算部 225 を制御するものである。

【0011】この相関係数値 $r(i)$ は、

$$\dots (6)$$

※し、又 m はマイクロホン 229-1, 229-2 間の距離とサンプリング周波数とによって決まる値で、

だエコー成分をエコーキャンセラによって打ち消すことができる。しかし、話者の発声とスピーカからの再生音とが同時にマイクロホンアレイに入力された場合、エコー成分(雑音)を打ち消す為のフィルタ係数の更新が収束しない状態となる。即ち、(4)式及び(5)式の残差信号 e は、減算器 216 によりエコー成分(雑音)を抑制できなかった成分と、話者の音声との和となるから、この残差信号 e を最小とするようにフィルタ係数の更新を行うと、目的音としての話者の音声もエコー成分(雑音)と共に抑制することになり、目的の雑音抑制ができない問題があった。

【0014】又音源位置検出と目的音強調との為の従来例に於いて、例えば、図 22 のマイクロホン 229-1, 229-2 の出力信号 $a(i), b(i)$ は、一般には、サンプル値の近傍に於いて自己相関がある。な

お、音源が白色雑音又はパルス雑音等の場合は、自己相関は小さくなり、音声等の場合は自己相関が大きくなる。前述の(6)式による相関関数値 $r(i)$ は、自己相関が大きい信号に対して、自己相関が小さい信号より i に対する値の変化が小さくなる。従って、正確な最大値を求めることが容易でなくなり、音源位置の検出を正確に且つ迅速に行うことが困難となる問題があった。

【0015】又目的音強調の為の同期加算を行う従来例に於いて、強調の程度はマイクロホンアレイを構成するマイクロホンの個数に依存することになり、目的音と雑音との間の相関が小さいと、 N 個のマイクロホンを用いることによりパワー比で N 倍の強調を行うことができるが、目的音と雑音との間の相関が大きいと、そのパワー比は小さくなる。従って、目的音と雑音との間の相関が大きい場合の目的音強調を行う為には、マイクロホンの個数を増加する必要がある、マイクロホンアレイが大型化する問題がある。又前述の(6)式による相関係数値から目的音の音源位置を検出する場合に、雑音等が大きい環境では、音源位置の検出が困難となる場合が多くなる。本発明の目的は、マイクロホンアレイを用いて雑音抑制、目的音強調、音源位置検出の処理を安定に且つ確実に行うことを可能とする。

【0016】

【課題を解決するための手段】本発明のマイクロホンアレイ装置は、(1)複数のマイクロホン $1-1 \sim 1-n$ を配列して構成したマイクロホンアレイを有するマイクロホンアレイ装置に於いて、マイクロホン $1-1 \sim 1-n$ の出力信号を入力するフィルタ $2-1 \sim 2-n$ と、マイクロホン $1-1 \sim 1-n$ の出力信号と、雑音源信号と、フィルタ $2-1 \sim 2-n$ を介したマイクロホン $1-1 \sim 1-n$ の出力信号の中の基準マイクロホン $1-1$ の出力信号から他のマイクロホン $1-2 \sim 1-n$ の出力信号を減算した残差信号とを入力し、この残差信号を基にした評価関数に従ってフィルタ $2-1 \sim 2-n$ の係数を求めるフィルタ係数計算部 4 とを備えている。

【0017】又(2)フィルタの前段に接続した遅延器と、複数のマイクロホンの出力信号と雑音源信号との相互相関関数値を求めて該相互相関関数値が最大値となる条件を基に前記遅延器の遅延量を求める遅延計算部とを設けることができる。従って、フィルタ係数計算部 4 には遅延器により位相がそろった信号が入力され、フィルタ係数の更新制御が容易となる。

【0018】又(3)雑音源信号を、スピーカを駆動する信号とする。即ち、マイクロホンアレイとスピーカとを有するシステムに於いて、スピーカからの再生音声がマイクロホンアレイに回り込んで雑音となるから、このスピーカを雑音源とした時、このスピーカを駆動する信号を雑音源信号として用いることにより、フィルタ係数計算部 4 に於ける処理が容易となる。

【0019】又(4)複数のマイクロホンからなるマイ

クロホンアレイと共に、雑音源信号を出力する補助マイクロホンを設けることができる。この場合は、マイクロホンアレイのみを有するシステムに於いて、補助マイクロホンの出力信号を雑音源信号として、フィルタ係数計算部 4 に於いてフィルタ係数の更新制御を行う。

【0020】又(5)フィルタ係数計算部に於けるフィルタ係数の更新処理に於ける畳み込み演算のフィルタのメモリ値に対して、重み付けを小さくする為の巡回型ローパスフィルタを設けることができる。

【0021】又(6)マイクロホンの出力信号を入力する線形予測フィルタと、マイクロホンの出力信号を入力して線形予測フィルタのフィルタ係数を線形予測分析に従って更新する線形予測分析部と、線形予測フィルタの出力信号の線形予測誤差信号を基に相関係数値を求め、この相関係数値が最大となる値を基に音源位置情報を出力する音源位置検出部とを備えることができる。

【0022】又(7)目的音源をスピーカとし、このスピーカを駆動する信号を入力して、複数のマイクロホン対応の線形予測フィルタに対するフィルタ係数更新の制御を行う線形予測分析部を設けることができる。この線形予測分析部は、マイクロホン対応の線形予測フィルタに対して共通化できることになる。

【0023】又(8)複数のマイクロホンの出力信号と音波の伝搬速度とを基に、マイクロホンの配置間隔に従って推定マイクロホンが配置されたと推定し、この推定マイクロホンの出力信号を、マイクロホンアレイを構成するマイクロホンの出力信号と共に出力する信号推定部と、この信号推定部のマイクロホンアレイを構成するマイクロホン及び推定マイクロホンの出力信号の位相を合わせて加算する同期加算部とを備えることができる。

【0024】又(9)マイクロホンアレイを構成する複数のマイクロホンの配置線上に、マイクロホンの配置間隔に従って配置した参照マイクロホンを設け、信号推定部は、マイクロホンアレイを構成する複数のマイクロホンの出力信号を基に、推定マイクロホンの配置位置及び推定マイクロホンの出力信号を補正する構成とすることができる。従って、推定マイクロホンの演算処理に於ける誤差を小さくして目的音強調を行うことができる。

【0025】又(10)参照マイクロホンの出力信号と信号推定部で推定した参照マイクロホンの配置位置の推定マイクロホンの出力信号との差の誤差信号に対して、聴覚特性に従った重み付けを行って聴覚感度が高い帯域の推定精度を高くする推定用係数決定部を設けることができる。

【0026】又(12)マイクロホンアレイに対する音源の方向を所定の角度に分割し、この分割した方向対応に、マイクロホンアレイを構成するマイクロホンの出力信号及びこの出力信号を基に推定した推定マイクロホンの出力信号とを出力する信号推定部と、この信号推定部の出力信号を位相を合わせて加算する同期加算部と、こ

の同期加算部の出力信号の最大値を基に音源位置情報を出力する音源位置検出部とを備えることができる。

【0027】又(12)複数のマイクロホンの出力信号を基に音源位置を検出する音源位置検出部と、音源を撮像するカメラと、このカメラの撮像信号を基に、音源位置を検出する検出部と、音源位置検出部からの位置情報と、検出部からの位置情報とを基に、音源の位置を示す音源位置情報を出力する統合判定処理部とを備えることができる。

【0028】

【発明の実施の形態】図1は本発明の第1の実施の形態の説明図であり、1-1~1-nはマイクロホンアレイを構成するn個のマイクロホン、2-1~2-nはフィルタ、3は加算器、4はフィルタ係数計算部、5は話者(目的音源)、6はスピーカ(雑音源)を示す。話者5からの音声信号がマイクロホン1-1~1-nに入力されて電気信号に変換され、フィルタ2-1~2-nと加算器3とを介して出力信号となり、ネットワーク等を介して相手側へ伝送される。又相手側からの音声信号を入力信号としてスピーカ6が駆動されて再生音声となる。それによって、話者5は、相手側と通話ができる。この場合、スピーカ6からの再生音声信号がマイクロホン1-1~1-nに入力されるから、話者5からの音声に対して雑音となる。従って、スピーカ6は目的音源に対して雑音源となる。

【0029】そこで、本発明に於いては、フィルタ係数*

$$c = [f_1(1) \cdots f_1(r)] \begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1r} \end{bmatrix}$$

$$- \sum_{i=2}^n [f_i(1) \cdots f_i(r)] \begin{bmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{ir} \end{bmatrix} \quad \cdots (9)$$

となる。この場合、加算器3に於いて、フィルタ2-1の出力信号から、フィルタ2-2~2-nの出力信号を減算することを示すものである。なお、 $f_1(1), f_1(2), \cdots, f_1(r), \cdots, f_i(1), f_i(2), \cdots, f_i(r)$ はフィルタのメモリの値を示す。

* 計算部4に、マイクロホン1-1~1-nの出力信号と、雑音源信号(雑音源としてのスピーカ6を駆動する為の入力信号)と、加算器3の出力信号(残差信号)とを入力して、フィルタ2-1~2-nの係数更新を行わせるものである。この場合、マイクロホン1-1を基準マイクロホンとして、フィルタ2-1の出力信号に対して、他のフィルタ2-2~2-nの出力信号を加算器3に於いて減算する構成としている。

【0030】又フィルタ2-1~2-nは、例えば、図2に示す構成とすることができる。同図に於いて、11-1~11-r-1は Z^{-1} の遅延器、12-1~12-rはフィルタ係数 $c_{p1}, c_{p2}, \cdots, c_{pr}$ を乗算する為の係数器、13, 14は加算器であり、rはフィルタの次数を示す。

【0031】雑音源(スピーカ6)からの信号を $x_p(i)$ とし、目的音源(話者5)からの信号を $y_p(i)$ とすると(但し、iはサンプル番号、pは1, 2, \cdots, n)、フィルタ2-1~2-nのメモリの値(フィルタへの入力信号と遅延器11-1~11-r-1の出力信号) $f_p(i)$ は、 $f_p(i) = x_p(i) + y_p(i) \quad \cdots (8)$ となる。

【0032】従来例のマイクロホンアレイを用いたエコーキャンセラでは、図1に於ける加算器3の出力信号eは、

【数3】

【0033】これに対して、本発明では、雑音源からの信号 $x_p(i)$ の位相を合わせてから畳み込むと、加算器3の出力信号 e' は、

【数4】

11

12

$$e' = [f_1(1)' \cdots f_1(r)'] \begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1r} \end{bmatrix} \\ - \sum_{i=2}^n [f_i(1)' \cdots f_i(r)'] \begin{bmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{ir} \end{bmatrix} \quad \cdots (10)$$

$$[f_p(1)' \cdots f_p(r)'] \\ = [x(1)(p) \cdots x(q)(p)] \begin{bmatrix} f_p(1) \cdots f_p(r) \\ f_p(2) \cdots f_p(r+1) \\ \vdots \\ f_p(q) \cdots f_p(q+r-1) \end{bmatrix} \quad \cdots (11)$$

となる。なお、 $x(1)(p), \cdots, x(q)(p)$ の (p) は、マイクロホン 1-1 ~ 1-n の位相を合わせた雑音源からの信号であることを示し、 q は畳み込み演算を行うサンプル数を示す。

【0034】 雑音源からの信号 $x_p(i)$ と目的音源か*

$$[f_p(1)' \cdots f_p(r)'] \\ = [x(1)(p) \cdots x(q)(p)] \begin{bmatrix} f_p(1) \cdots f_p(r) \\ f_p(2) \cdots f_p(r+1) \\ \vdots \\ f_p(q) \cdots f_p(q+r-1) \end{bmatrix} \\ = [x(1)(p) \cdots x(q)(p)] \begin{bmatrix} \{x_p(1)+y_p(1)\} \cdots \{x_p(r)+y_p(r)\} \\ \{x_p(2)+y_p(2)\} \cdots \{x_p(r+1)+y_p(r+1)\} \\ \vdots \\ \{x_p(q)+y_p(q)\} \cdots \\ \{x_p(q+r-1)+y_p(q+r-1)\} \end{bmatrix} \\ \doteq \left[\sum_{i=1}^q x(i)(p) * x_p(i) \cdots \sum_{i=1}^q x(i)(q) * x_p(r+i-1) \right] \quad \cdots (12)$$

となる。

【0035】 この(12)式から判るように、 $[f_p(1)', \cdots, f_p(r)']$ に於ける目的音源からの信号 $v_p(i)$ の影響が小さくなる。この(12)式を用いて(10)式の信号 e' を求め、これを基に評価関数 $J = (e')^2$ を求め、この評価関数 $J = (e')^2$ を基にフィルタ 2-1 ~ 2-n のフィルタ係数の更新制御を行うものである。即ち、話者(目的音源) 5 とスピーカ(雑音源) 6 とから同時にマイクロホン 1-1 ~ 1-n に音声が入力される状態に於いても、マイクロホン 1-1 ~ 1-n の出力信号に含まれる雑音源信号は、フィルタ係数計算部 4 に入力されるスピーカ 6 を駆動する為の入力信号との相関が大きく、又目的音源信号との相関は小さくなるから、評価関数 $J = (e')^2$ に従ったフィルタ係数の更新制御が可能となる。従って、加算器 3 の出力信号は、雑音が抑制された話者 5 の音声信号とな

*らの信号 $v_p(i)$ との両方が同時に入力される場合、即ち、話者 5 の発声とスピーカ 6 からの再生音声とが同時に生じた場合、異なる人間の音声であるから両者の相関は小さいものであり、従って、(11)式は、

【数 5】

る。

【0036】 図 3 は本発明の第 2 の実施の形態の説明図であり、図 1 と同一符号は同一部分を示し、8-1 ~ 8-n は遅延器 ($Z^{-d_1} \sim Z^{-d_n}$)、9 は遅延計算部である。この実施の形態は、マイクロホン 1-1 ~ 1-n からの信号の位相を合わせるように遅延器 8-1 ~ 8-n の遅延サンプル数を遅延計算部 9 に於いて算出し、フィルタ 2-1 ~ 2-n のフィルタ係数をフィルタ係数計算部 4 に於いて算出して更新制御する。その為、遅延計算部 9 に、マイクロホン 1-1 ~ 1-n の出力信号と、スピーカ 6 を駆動する為の入力信号(雑音源信号)とを入力し、フィルタ係数計算部 4 に、遅延器 8-1 ~ 8-n の出力信号と、加算器 3 の出力信号と、スピーカ 6 を駆動する為の入力信号(雑音源信号)とを入力する。

【0037】 マイクロホン 1-1 ~ 1-n の出力信号を $g_p(i)$ (但し、 $p=1, 2, \cdots, n$ 、 i =サン

13

ル番号)とし、雑音源からの信号 $x(i)$ との相互相関 * 関数値 $R_p(i)$ を次式に示すように求める。

$$R_p(i) = \sum_{j=1}^s x_p(i+j) * x(i) \quad \dots (13)$$

なお、 $\sum_{j=1}^s$ は $i=1$ から $i=s$ までの加算を示し、 s は畳み込み演算を行うサンプル数を示す。このサンプル数 s は通常は数 10 ～ 数 100 サンプルとすることができる。又雑音源からマイクロホンまでの距離に対応した最大遅延サンプル数を D とすると、(13) 式に於て※

$$D = (\text{サンプリング周波数}) * (\text{雑音源とマイクロホンとの間の最大距離}) / (\text{音速})$$

$$= 8000 * (50 / 34000) = 11.76 \dots \approx 12$$

とすることができる。従って、この場合の i は、 $i=1 \sim 12$ の範囲とする。又雑音源とマイクロホンとの間の距離の最大値を 1 m とすると、最大遅延サンプル数 D は 24 となる。

【0039】又 (13) 式により求めた相互相関関数値 $R_p(i)$ の絶対値が最大となる時の i の値 i_p ($p=1, 2, \dots, n$) を求め、更に、 i_p の最大値 i_{max} を求める。この処理は、図 4 の (A1) ～ (A11) に示すステップに従ったものとなる。即ち、 i_{max} = 初期値 (例えば、0) とし、且つ $p=1$ とし (A1)、次に、 R_{pmax} = 初期値 (例えば、0.0)、 i_p = 初期値 (例えば、0) とし、且つ $i=0$ とし (A2)、前述の (13) 式による相互相関関数値 $R_p(i)$ を求める (A3)。

【0040】そして、相互相関関数値 $R_p(i)$ が R_{pmax} より大きいかな否かを判定し (A4)、大きい場合★

$$d_p = i_{max} - i_p$$

それによって、遅延器 8-1 ～ 8-n の遅延サンプル数 $d_1 \sim d_n$ が遅延計算部 9 によって設定される。

【0042】又フィルタ 2-1 ～ 2-n は、前述のよう★30

$$out_p = \sum_{i=1}^n c_{pi} * f_p(i) \quad \dots (15)$$

となる。なお、 $\sum_{i=1}^n$ は、 $i=1$ から $i=n$ までの加算を示し、 c_{pi} はフィルタ係数、 $f_p(i)$ はフィルタのメモリの値を表し、この場合のフィルタの入力信号でもある。

【0043】又フィルタ係数計算部 4 に於いては、現在◆

$$f_p(i)' = \sum_{j=1}^q x(i) * f_p(i+j-1) \quad \dots (16)$$

となる。なお、 $\sum_{j=1}^q$ は $i=1$ から $i=q$ までの加算を示し、 q は相互相関関数値を計算する時の畳み込み演算を行うサンプル数を示し、一般には、数 10 ～ 数 10 * 40

$$e' = \sum_{i=1}^n [f_{1i}(i)' * c_{1i}] - \sum_{i=1}^n \sum_{j=1}^n [f_{ji}(i)' * c_{ji}] \quad \dots (17)$$

として求めることができ、畳み込み演算であるから、デジタル・シグナル・プロセッサ (DSP) によって算出することができる。この場合、加算器 3 は、フィルタ 2-1 を介した基準マイクロホン 1-1 の出力信号から、フィルタ 2-2 ～ 2-n を介した他のマイクロホン 1-2 ～ 1-n の出力信号を減算して出力信号 e' を出力することになる。

【0045】前述の加算器 3 の出力信号 e' を誤差信号 50

14

* 関数値 $R_p(i)$ を次式に示すように求める。

※ i は、 $i=0, 1, 2, \dots, D$ とする。

【0038】例えば、雑音源とマイクロホンとの間の距離の最大値を 50 cm、サンプリング周波数を 8 kHz とすると、音速は約 340 m/s であるから、最大遅延サンプル数 D は、

★は、その時の $R_p(i)$ を R_{pmax} とし (A5)、小さい時は、 $i=i+1$ とする (A6)。そして、 $i \leq D$ か否かを判定し (A7)、 i が最大遅延サンプル数 D 以下の時はステップ (A3) に移行し、 i が最大遅延サンプル数 D を超えるとステップ (A8) に移行する。このステップ (A8) に於いて、 i_p が i_{max} より大きいかな否かを判定し、大きい場合は、その時の i_p を i_{max} とし (A9)、大きくない場合は、 $p=p+1$ とし (A10)、 $p \leq n$ か否かを判定し (A11)、 $p \leq n$ の条件の時はステップ (A2) に移行し、その条件を満足しない時は、相互相関関数値 $R_p(i)$ の検索が終了し、 $i \leq D$ の範囲に於ける i_p の最大値 i_{max} が得られる。

【0041】前述の最大値検出により得られた i_p と i_{max} とを用いて、遅延器の遅延サンプル数 d_p を次式によって求める。

$$\dots (14)$$

☆に、図 2 に示す構成を適用することができるものであり、各フィルタ 2-1 ～ 2-n の出力信号を out_p ($p=1, 2, \dots, n$) とすると、

◆と過去とのフィルタ 2-1 ～ 2-n の入力信号と、雑音源からの信号との相互相関関数値を計算して、フィルタ係数の更新を行うものであり、相互相関関数値 $f_p(i)'$ は、

【0044】このような相互相関関数値 $f_p(i)'$ を用いて、加算器 3 の出力信号 e' を求める。即ち、

として、評価関数 $J = (e')^2$ とするもので、この評価関数 $J = (e')^2$ を基にフィルタ係数を求める。例えば、前述のように、最急降下法により求めることができ、次式によりフィルタ係数 $c_{11}, c_{12}, \dots, c_{1r}, \dots, c_{n1}, c_{n2}, \dots, c_{nr}$ を求めることができる。

【数 6】

15

$$\begin{bmatrix} c11 \\ c12 \\ \vdots \\ c1r \end{bmatrix} = \begin{bmatrix} c11_{old} \\ c12_{old} \\ \vdots \\ c1r_{old} \end{bmatrix} - t1 * \begin{bmatrix} n(1)' \\ n(2)' \\ \vdots \\ n(r)' \end{bmatrix}$$

... (18)

$$t1 = \alpha * (e' / f1_{norm})$$

$$\begin{bmatrix} cp1 \\ cp2 \\ \vdots \\ cpr \end{bmatrix} = \begin{bmatrix} cp1_{old} \\ cp2_{old} \\ \vdots \\ cpr_{old} \end{bmatrix} + tp * \begin{bmatrix} fp(1)' \\ fp(2)' \\ \vdots \\ fp(r)' \end{bmatrix}$$

... (19)

$$tp = \alpha * (e' / fp_{norm})$$

$$p = 2, 3, \dots, n$$

なお、ノルム $f_{p_{norm}}$ は、(3) 式に対応したものであり、

$$f_{p_{norm}} = [(f_p(1)')^2 + (f_p(2)')^2 + \dots + (f_p(r)')^2]^{1/2}$$

... (20)

となる。又 (18)、(19) 式に於ける α は、前述のように、定数で、フィルタ係数の最適値への収束の速さ

$$e' = out1 - \sum_{i=1}^n out_i$$

となり、フィルタ 2-1 ~ 2-n への入力信号の位相を遅延器 8-1 ~ 8-n によって揃えることができるから、フィルタ係数計算部 4 によるフィルタ係数の更新が容易となり、且つ話者 5 とスピーカ 6 とから同時に発音する状態の場合に於いても、フィルタ係数の更新制御が可能となり、雑音源としてのスピーカ 6 からマイクロホン 1-1 ~ 1-n に回り込む雑音を確実に抑制することができる。

【0047】図 5 は本発明の第 3 の実施の形態の説明図であり、図 1 と同一符号は同一部分を示し、16 は雑音源、21 は補助マイクロホンである。この補助マイクロホン 21 は、マイクロホンアレイを構成するマイクロホン 1-1 ~ 1-n と同一構成のマイクロホンとすることができる。

【0048】この実施の形態は、図 1 に示す実施の形態とほぼ同一であるが、補助マイクロホン 21 の出力信号を、雑音源の信号としてフィルタ係数計算部 4 に入力するものである。従って、雑音源 16 は、話者 5 又は任意の目的音源に対して、スピーカ以外の空調音等の任意の雑音源とした場合に於いても、図 1 に関連して説明した★40

$$f_p(i)' = \beta * f_p(i)'_{old} + (1-\beta) * [x(1) * f_p(i)'] \quad \dots (22)$$

なお、係数 β は、 $0.0 < \beta < 1.0$ の条件で設定され、又 $f_p(i)'_{old}$ はローパスフィルタのメモリ（遅延器 25）の値を示す。

【0051】この巡回型ローパスフィルタを用いることにより、過去の信号の重み付けを小さくして、畳み込み演算に於ける出力値が過大になることを防止し、安定に相互相関関数値 $f_p(i)'$ を求めることができる。

【0052】図 8 は DSP（デジタル・シグナル・プ

※と精度とを表すものである。

【0046】従って、加算器 3 の出力信号 e' は、

... (21)

★ように、フィルタ係数の更新に用いる評価関数 $J = (e')^2$ を基に雑音抑制が可能となる。

【0049】図 6 は本発明の第 4 の実施の形態の説明図であり、図 3 及び図 5 と同一符号は同一部分を示す。この実施の形態は、図 3 に示す実施の形態とほぼ同一であるが、補助マイクロホン 21 の出力信号を、雑音源の信号として遅延計算部 9 及びフィルタ係数計算部 4 に入力するものである。従って、図 3 に示す実施の形態の場合と同様に、遅延計算部 9 により遅延器 8-1 ~ 8-n の遅延サンプル数を制御し、フィルタ係数計算部 4 によりフィルタ 2-1 ~ 2-n のフィルタ係数の更新制御を行って、雑音抑制を行うことができる。

【0050】図 7 は本発明の実施の形態のフィルタ係数更新処理に於けるローパスフィルタの説明図であり、22、23 は係数器、24 は加算器、25 は遅延器である。前述の相互相関関数値 $f_p(i)'$ を、図 7 に示すローパスフィルタを用いて算出する場合を示し、係数器 23 の係数を β とし、係数器 22 の係数を $1-\beta$ とした場合を示す。

ロセッサ）を用いた本発明の実施の形態の説明図であり、1-1 ~ 1-n はマイクロホンアレイを構成するマイクロホン、30 はデジタル・シグナル・プロセッサ（DSP）、31-1 ~ 31-n はローパスフィルタ（LPF）、32-1 ~ 32-n は AD 変換器（A/D）、33 は DA 変換器（D/A）、34 はローパスフィルタ（LPF）、35 は増幅器、36 はスピーカを示す。

【0053】図1に示す実施の形態に於けるフィルタ2-1~2-nとフィルタ係数計算部4及び図3に示す実施の形態に於けるフィルタ2-1~2-nとフィルタ係数計算部4及び遅延器8-1~8-nと遅延計算部9とは、繰り返し処理と積和演算と条件分岐との組合せによって実現できるから、このような処理をデジタル・シグナル・プロセッサ30の演算機能によって実現するものである。

【0054】又ローパスフィルタ31-1~31-n、34は、例えば、音声帯域以外の信号成分を除去するものであり、又AD変換器32-1~32-nは、マイクロホン1-1~1-nの出力信号をローパスフィルタ31-1~31-nを介して入力してデジタル信号に変換するもので、例えば、8kHzでサンプリングして、デジタル・シグナル・プロセッサ30に於いて処理するビット数に対応して8ビットや14ビット等に変換する。

【0055】又ネットワーク等を介した入力信号は、DA変換器33によりアナログ信号に変換され、ローパスフィルタ34を介して増幅器35に入力され、増幅してスピーカ36を駆動することになる。この場合のスピーカ36の再生音は、マイクロホン1-1~1-nに対しては雑音となる。しかし、前述のように、デジタル・シグナル・プロセッサ30によるフィルタ係数の更新等によって雑音を抑制することができる。

【0056】図9は本発明の実施の形態のDSP（デジタル・シグナル・プロセッサ）の処理機能の説明図であり、図3及び図8と同一符号は同一部分を示し、図8に於けるローパスフィルタ31-1~31-n、34と、AD変換器32-1~32-nと、DA変換器33と、増幅器35との図示を省略している。又フィルタ係数計算部4は、相互相関計算部41とフィルタ係数更新部42とを含み、遅延計算部9は、相互相関計算部43と最大値検出部44と遅延サンプル数計算部45とを含むものである。

【0057】遅延計算部9の相互相関計算部43は、マイクロホン1-1~1-nの出力信号 $g_p(i)$ と、雑音源としてのスピーカ36の駆動信号とを入力し、(13)式に示す相互相関関数値 $R_p(i)$ を算出する。又最大値検出部44は、図4に示すフローチャートに従って相互相関関数値 $R_p(i)$ の最大値を検出し、遅延サンプル数計算部45は、最大値検出により得られた i_p と i_{max} とを用いて、遅延器8-1~8-nの遅延サンプル数 d_p を(14)式に従って求め、遅延器8-1~8-nの遅延サンプル数を設定する。

【0058】又フィルタ係数計算部4の相互相関計算部41は、遅延器8-1~8-nによって雑音源の信号の位相を合わせた信号と、雑音源としてのスピーカ36の駆動信号と、加算器3の出力信号とを入力し、前述の

(16)式に従って相互相関関数値 $f_p(i)'$ を算出

する。この相互相関関数値 $f_p(i)'$ の算出過程に於いて、図7に示すローパスフィルタの処理を含めることができる。又フィルタ係数更新部42は、(17)、

(18)、(19)式に従ってフィルタ係数 c_{pr} を算出し、例えば、図2に示す機能のフィルタ2-1~2-nのフィルタ係数の更新を行うものである。

【0059】図10は遅延器の説明図であり、46はメモリ、47は書込制御部、48は読出制御部、9は遅延計算部である。デジタル・シグナル・プロセッサの内部メモリを用いて遅延器を実現した場合を示し、メモリ46は、遅延サンプル数の最大値Dの領域を有し、書込制御部47の制御によって書込みが行われ、又読出制御部48の制御により読出される。又遅延計算部9により算出された遅延サンプル数 d_p の間隔で書込ポインタWPと読出ポインタRPとが設定され、且つ点線矢印方向に、書込み読出しのタイミング毎にシフトされる。従って、書込ポインタWPにより指示されたアドレスに書込まれた信号は、設定された遅延サンプル数 d_p の後の読出ポインタRPにより指示された時に読出される。

【0060】図11は本発明の第5の実施の形態の説明図であり、51-1、51-2はマイクロホンアレイを構成するマイクロホン、52-1、52-2は線形予測フィルタ、53-1、53-2は線形予測分析部、54は音源位置検出部、55は話者等の音源を示す。マイクロホンアレイを構成するマイクロホンは更に多数設けることも可能であるが、以下説明の便宜上2個のマイクロホン51-1、51-2を設けた場合について説明する。

【0061】マイクロホン51-1、51-2の出力信号 $a(i)$ 、 $b(i)$ をそれぞれ線形予測分析部53-1、53-2と、線形予測フィルタ52-1、52-2とに入力し、線形予測分析部53-1、53-2に於いて自己相関関数値を求めて線形予測係数を算出し、この線形予測係数を用いて線形予測フィルタ52-1、52-2のフィルタ係数の更新を行い、線形予測フィルタ52-1、52-2の出力信号の線形予測誤差信号を基に、音源検出部54に於いて音源55の位置を検出し、音源位置情報を出力する。

【0062】図12は図11に示す各部の機能を更に詳細に示すもので、図11と同一符号は同一部分を示し、56-1、56-2は自己相関関数値計算部、57-1、57-2は線形予測係数計算部、58は相関係数値計算部、59は位置検出処理部である。線形予測分析部53-1、53-2は、自己相関関数値計算部56-1、56-2と、線形予測係数計算部57-1、57-2とを含む構成であり、マイクロホン51-1、51-2の出力信号 $a(i)$ 、 $b(i)$ が自己相関関数値計算部56-1、56-2に入力される。

【0063】線形予測分析部53-1の自己相関関数値計算部56-1は、マイクロホン51-1の出力信号 a

10

20

30

40

50

(i) を基に自己相関関数値 $R a(i)$ を次式により算 * * 出する。

$$R a(i) = \sum_{j=1}^n a(i) * a(i+j) \quad \dots (23)$$

なお、 $\sum_{j=1}^n$ は $j=1$ から $j=n$ までの加算を示し、 n は畳み込み演算のサンプル数で、一般に数 100 の値となる。又 q を線形予測フィルタの次数とすると、 $0 \leq i \leq q$ となる。

【0064】又線形予測係数計算部 57-1 は、自己相関関数値 $R a(i)$ を基に線形予測係数 $\alpha a_1, \alpha a_2, \dots, \alpha a_q$ を算出する。この線形予測係数は、相関法、偏自己相関法、共分散法等の既に知られた各種の方法によって求めることができる。従って、前述のデジタル・シグナル・プロセッサ (DSP) の演算機能によっても実現できる。

【0065】又マイクロホン 51-2 に対応する線形予測分析部 53-2 に於いても、自己相関関数値計算部 5 *

$$q = (\text{サンプリング周波数}) * (\text{マイクロホン間距離}) / (\text{音速})$$

... (24)

により定まる値であり、右辺は前述の (7) 式と同様のものとなる。

【0067】又音源位置検出部 54 は、相関係数値計算部 58 と、位置検出処理部 59 とを含み、相関係数値計算部 58 は、線形予測フィルタ 52-1, 52-2 の出力信号、即ち、マイクロホン 51-1, 51-2 の出力信号 $a(i), b(i)$ の線形予測誤差信号 a'

$(i), b'(i)$ を用いて相関係数値 $r'(i)$ を算出する。この場合の i の範囲は、 $-q \leq i \leq q$ となる。

【0068】位置検出処理部 59 は、相関係数値 $r'(i)$ の値を最大とする i の値 i_{max} を求め、その値 i_{max} により音源 55 の位置を示す音源位置情報を出力する。この場合の音源位置と i_{max} との関係は図 13 に示すものとなる。即ち、 $i_{max} = 0$ の場合は、音源 55 がマイクロホン 51-1, 51-2 の前方又は後方で、マイクロホン 51-1, 51-2 から等距離の位置に存在することになる。又 $i_{max} = q$ の場合は、マイクロホン 51-1, 51-2 の配置線上のマイクロホン 51-1 側に存在し、 $i_{max} = -q$ の場合は、マイクロホン 51-2 側に存在することになる。なお、マイクロホンを 3 個以上とすれば、音源までの距離を含めて音源位置を検出することができる。

【0069】音声信号は、自己相関関数値が一般に大きいものであり、マイクロホン 51-1, 51-2 の出力信号 $a(i), b(i)$ を用いて相関係数値 $r(i)$ を求める従来例は、相関係数値 $r(i)$ の i に対する値の変化が小さいことにより、音源位置の検出が容易でないものであったが、前述の本発明の実施の形態によれば、自己相関関数値が大きい場合でも、線形予測誤差信号を用いて相関係数値 $r'(i)$ を求めるもので、等価的に、自己相関を小さくすることに相当し、音源位置の検出が容易となる。

【0070】図 14 は本発明の第 6 の実施の形態の説明

※ 6-2 により、マイクロホン 51-2 の出力信号 b

(i) を基に自己相関関数値 $R b(i)$ を (23) 式と同様にして算出し、線形予測係数計算部 57-2 により、線形予測係数 $\alpha b_1, \alpha b_2, \dots, \alpha b_q$ を算出する。

【0066】線形予測フィルタ 52-1, 52-2 は、 q 次の FIR フィルタの構成とすることができるもので、フィルタ係数 c_1, c_2, \dots, c_q は、それぞれ線形予測係数 $\alpha a_1, \alpha a_2, \dots, \alpha a_q, \alpha b_1, \alpha b_2, \dots, \alpha b_q$ によって更新される。この線形予測フィルタ 52-1, 52-2 のフィルタ次数 q は、

図であり、図 11 と同一符号は同一部分を示し、53A は線形予測分析部、55A は音源としてのスピーカである。音源としてのスピーカ 55A の駆動信号を線形予測分析部 53A に入力することにより、音源の信号を線形予測分析し、線形予測係数を求めるもので、線形予測フィルタ 52-1, 52-2 に対して共通の線形予測分析部 53A として、マイクロホン 51-1, 51-2 の出力信号 $a(i), b(i)$ の線形予測誤差信号を求め、音源位置検出部 54 は、その線形予測誤差信号を用いて相関係数値 $r'(i)$ を求め、音源の位置を検出することができる。

【0071】図 15 は本発明の第 7 の実施の形態の説明図であり、61-1, 61-2 はマイクロホンアレイを構成するマイクロホン、62 は信号推定部、63 は同期加算部、65 は音源を示し、例えば、2 個のマイクロホン 61-1, 61-2 の配置線上に推定位置として点線で示すマイクロホン 64-1, 64-2, ... が存在しているものとして同期加算部 63 に於いてマイクロホン 61-1, 61-2 の出力信号の同期加算を行って目的音強調を行う構成を示すものである。

【0072】図 16 は本発明の第 7 の実施の形態の機能ブロック図であり、図 15 と同一符号は同一部分を示し、66 は粒子速度計算部、67 は推定処理部、68-1, 68-2, ... は遅延器、69 は加算器である。マイクロホンアレイを構成する 2 個のマイクロホン 61-1, 61-2 の配置線に対して音源 65 が θ の方向に位置している場合を示し、又マイクロホン 61-1, 61-2 の配置線上に沿って点線で示すマイクロホン 64-1, 64-2, ... が配置されていると推定して処理するものである。

【0073】又信号推定部 62 は、粒子速度計算部 66 と、推定処理部 67 とを含む構成を有する。又音源 65 からの音波は、その伝搬を波動方程式 (Wave Equatio

n) で表すことができる。この場合、音圧を P、粒子速度を V、媒質の体積弾性率を K、媒質の密度を ρ とする *

$$-\partial V / \partial x = (1/K) (\partial P / \partial t)$$

$$-\partial P / \partial t = \rho (\partial V / \partial t)$$

… (25)

の関係で表されることが知られている。

※ (i, 1) とし、その音圧差によって粒子速度 V を求め

【0074】粒子速度計算部 66 は、マイクロホン 61-1 の出力信号 a (i) の振幅を音圧 P (i, 0)、マイクロホン 61-2 の出力信号 b (i) の振幅を音圧 P ※

る。即ち、マイクロホン 61-1 に於ける粒子速度 V (i+1, 0) は、

$$V (i+1, 0) = V (i, 0) + [P (i, 1) - P (i, 0)]$$

… (26)

と表すことができる。なお、i はサンプル番号である。★と、

【0075】推定処理部 67 は、推定位置を x とする ★

$$P (i, x+1) = P (i, x)$$

$$+ \beta (x) [V (i+1, x) - V (i, x)]$$

$$V (i+1, x) = V (i+1, x-1)$$

$$+ [P (i, x-1) - \rho (i, x)] \quad \dots (27)$$

によりマイクロホン 64-1, 64-2, … の推定位置を求めることができる。なお、 $\beta (x)$ は推定係数である。

☆力する。従って、2 個のマイクロホン 61-1, 61-2 からなるマイクロホンアレイでもって、更に多数のマイクロホンを配置したマイクロホンアレイと同様に同期加算による目的音強調が可能となる。

【0076】従って、マイクロホン 61-2 の配置位置を $x=1$ 、マイクロホン 61-1 の配置位置を $x=0$ とすると、推定位置のマイクロホン 64-1 は $x=2$ 、マイクロホン 64-2 は $x=3$ となり、推定処理部 62 は、2 個のマイクロホン 61-1, 61-2 を用いて、マイクロホン 64-1, 64-2, … が恰も配置されているかのように、マイクロホン 64-1, 64-2, … のそれぞれの出力信号を同期加算部 63 に入☆

【0077】同期加算部 63 は、遅延器 68-1, 68-2, … と、加算器 69 とを含み、遅延サンプル数を d とすると、遅延器 68-1, 68-2, … は、 Z^{-d} , Z^{-2d} , Z^{-3d} , … とし、マイクロホン 61-1, 61-2 の配置線に対する音源位置を前述の実施の形態によって求めた角度 θ を基に、

$$d = (\text{サンプリング周波数}) * (\text{マイクロホン間距離}) * \cos \theta / (\text{音速})$$

… (28)

によって遅延サンプル数 d を求める。

【0078】それによって、マイクロホン 61-1, 61-2 と推定位置のマイクロホン 64-1, 64-2, … のそれぞれの出力信号を遅延器 68-1, 68-2, … によって位相を合わせ、加算器 69 により加算して、同期加算による目的音強調処理を行うことができる。従って、少ない個数のマイクロホンを用いて、目的音を推定マイクロホン数に対応したパワーになるように強調できる。

【0079】図 17 は本発明の第 8 の実施の形態の説明図であり、図 15 と同一符号は同一部分を示し、71 は参照マイクロホン、72 は減算器、73 は重み付けフィ

30◆ ルタ、74 は推定用係数決定部である。この実施の形態は、位置 $x=0$ のマイクロホン 61-1 と、位置 $x=1$ のマイクロホン 61-2 との間隔と同一間隔で位置 $x=2$ に参照マイクロホン 71 を配置し、推定位置誤差を減算器 72 により求め、重み付けフィルタ 73 により聴覚特性を与えて、推定用係数決定部 74 により推定係数 $\alpha (x)$, $\beta (x)$, $\gamma (x)$ を決定する場合を示す。

【0080】即ち、位置 $x=2$ のマイクロホン 64-1 (参照マイクロホン 71 の位置の推定マイクロホン) の推定信号 P (i, 2) と参照マイクロホン 71 の出力信号 ref (i) との差の推定誤差 e (i) を減算器 72 により求める。

$$e (i) = P (i, 2) - \text{ref} (i)$$

$$= P (i, 1) + \beta (2) [V (i+1, 1) - V (i, 1)]$$

$$- \text{ref} (i)$$

… (29)

【0081】この推定誤差 e (i) の平均パワーが最小となるように、推定用係数決定部 74 に於いて推定係数 $\beta (2)$ を決定することができる。即ち、信号推定部 62 (図 15 又は図 16 参照) は、この推定誤差 e (i) の平均パワーを最小とする推定係数 $\beta (2)$ を $x=2$, 3, 4, … を用いて、推定マイクロホン 64-1, 50

64-2, … の出力信号を推定処理して出力することができる。

【0082】又図 17 に於いては、重み付けフィルタ 73 により、推定誤差 e (i) に聴覚特性に従った重み付けを行うもので、聴覚特性は、等ラウドネス曲線として知られているように、4 kHz 近傍の感度が高いを示し

ている。そこで、推定誤差 $e(i)$ に対して感度の高い 4kHz 近傍の帯域に対して重み付けを大きくするものである。従って、位置 $x=2$ 以降の推定マイクロホンの出力信号の処理に於いても、聴覚の感度が大きい帯域の推定誤差を小さくして、同期加算により目的音強調を行うことができる。

【0083】図18は本発明の第9の実施の形態の説明図であり、61-1、61-2はマイクロホンアレイを構成するマイクロホン、62-1、62-2、...62-sは信号推定部、63-1、63-2、...63-sは同期加算部、64-1、64-2、...は推定マイクロホン、65は音源、80は音源位置検出部である。

【0084】マイクロホン61-1、61-2からなるマイクロホンアレイの方向に対して角度 $\theta_0, \theta_1, \dots, \theta_s$ に分割し、それぞれ分割した角度 $\theta_0, \theta_1, \dots, \theta_s$ に対応し、信号推定部62-1~62-sと同期加算部63-1~63-sとを設ける。各信号推定部62-1~62-sは、予め推定係数 $\beta(x, \theta)$ を求めておくもので、例えば、図17に示すように、参照マイクロホンを設けて、推定係数 $\beta(x, \theta)$ を設定する。

【0085】同期加算部63-1~63-sは、信号推定部62-1~62-sの出力信号の位相を合わせて加算するものであり、それぞれ角度 $\theta_0, \dots, \theta_s$ の方向に対応した出力信号を得ることができる。そこで、音源位置検出部80は、各同期加算部63-1~63-sの出力信号のパワーを比較し、最大値のパワーの出力信号対応の角度を音源65の方向と判定し、音源位置情報を出力する。又最大値のパワーの出力信号を目的音強調信号として出力することができる。

【0086】図19は本発明の第10の実施の形態の説明図であり、90はテレビカメラ等のカメラ、91-1~91-2はマイクロホンアレイを構成するマイクロホン、92は音源位置検出部、93は音源の位置を検出する検出部としての顔位置検出部、94は統合判定処理部、95は音源を示す。

【0087】マイクロホン91-1、91-2と音源位置検出部92とは、前述の各実施の形態の何れかを適用した構成として、音源位置検出部92から音源95の位置情報を統合判定処理部94に入力する。又テレビカメラやデジタルカメラ等のカメラ90により話者を撮像し、話者の顔の位置を検出する。例えば、顔のテンプレートをを用いたテンプレートマッチング法により顔の位置を検出する方式や、カラー映像信号を基に肌色の領域を抽出して、顔の位置を検出する方式等を適用することができる。又統合判定処理部94は、音源位置検出部92による位置情報と、顔位置検出部93による位置検出情報とを基に、音源95の位置を判定して音源位置情報を出力する。

【0088】例えば、マイクロホン91-1、91-2の配置線とカメラ90の撮像方向とに対して、話者（音源）の方向を複数の角度 $\theta_0, \dots, \theta_s$ に分割し、マイクロホン91-1、91-2の出力信号の線形予測誤差を用いた相関係数値算出による音源位置検出、或いは、マイクロホン91-1、91-2とその配置線上の推定マイクロホンとの出力信号を用いた音源位置検出等の手段により、音源方向の確率を示す位置情報 $\text{inf-A}(\theta)$ を求める。又カメラ90からの映像信号を用いた話者（音源）の顔の方向の確率を示す位置情報 $\text{inf-V}(\theta)$ とを求める。そして、統合判定処理部94は、それぞれの位置情報 $\text{inf-A}(\theta)$ 、 $\text{inf-V}(\theta)$ との積 $\text{res}(\theta)$ を算出し、この積 $\text{res}(\theta)$ が最大となる角度 θ を音源位置情報として出力する。従って、音源95の方向を一層正確に検出することができる。又音源95の方向を検出してカメラ90のズーム等自動制御によって、音源95を拡大撮像することも可能となる。

【0089】本発明は、前述の各実施の形態のみに限定されるものではなく、種々付加変更することができるものであり、雑音抑制、目的音強調、音源位置検出等の目的に対応して前述の実施の形態を組み合わせることができる。又目的音強調や音源位置検出は、話者等の音声についてのみでなく、他の音波を発する音源の検出等荷も適用可能である。

【0090】

【発明の効果】以上説明したように、本発明は、マイクロホンアレイを構成するマイクロホン1-1~1-nの出力信号と、スピーカ6の駆動信号、補助マイクロホンの出力信号等の雑音源信号と、加算器3の出力の残差信号とをフィルタ係数計算部4に入力して、マイクロホン1-1~1-nの出力信号を入力するフィルタ2-1~2-nのフィルタ係数の更新制御を行うことにより雑音抑制を行うもので、目的音としての話者の音声と、雑音としての音声とが同時にマイクロホン1-1~1-nに入力される場合でも、両者の相互相関係数値が小さいことから、目的音としての話者の音声による影響を低減して、フィルタ係数の更新制御を継続して雑音抑制を行うことができる。

【0091】又フィルタ2-1~2-nの前段に遅延器を接続して、雑音信号の位相を合わせることにより、フィルタ2-1~2-nのフィルタ係数の更新制御が容易となるから、目的音としての話者の音声と、雑音としての音声等とが同時にマイクロホン1-1~1-nに入力された場合でも、雑音抑制が容易となる。

【0092】又マイクロホンアレイの出力信号又は目的音源の信号を入力して線形予測分析を行って、マイクロホンの出力信号を入力する線形予測フィルタのフィルタ係数を更新し、線形予測フィルタの出力信号を基に音源位置を検出することにより、目的音源の話者の音声と、

雑音源からの音声とが同時にマイクロホンに入力される場合でも、線形予測分析により音声信号の近傍サンプルの自己相関関数値を小さくして、目的音源の位置を確実に検出することができる。従って、その目的音源からの音声の強調、或いは、目的音源の音声以外を雑音として抑圧することができる。

【0093】又マイクロホンアレイを構成するマイクロホンの配置間隔に従った間隔の推定マイクロホンの出力信号も含めて同期加算を行うことにより、少ない個数のマイクロホンで、多数のマイクロホンを用いたマイクロ

ホンアレイと同様な目的音強調及び目的音源位置の検出を行うことができる利点がある。

【0094】又マイクロホンアレイによる音源位置の検出と目的音源の撮像信号による位置検出とを統合して判定することにより、目的音源の位置を迅速に且つ正確に検出することができる。

【図面の簡単な説明】

【図 1】 本発明の第 1 の実施の形態の説明図である。

【図 2】 フィルタの説明図である。

【図 3】 本発明の第 2 の実施の形態の説明図である。

【図 4】 本発明の第 2 の実施の形態に於ける遅延計算部の処理フローチャートである。

【図 5】 本発明の第 3 の実施の形態の説明図である。

【図 6】 本発明の第 4 の実施の形態の説明図である。

【図 7】 本発明の実施の形態のフィルタ係数更新処理に於けるローパスフィルタの説明図である。

【図 8】 D S P を用いた本発明の実施の形態の説明図である。

10

20

*

* 【図 9】 本発明の実施の形態の D S P の処理機能の説明図である。

【図 10】 遅延器の説明図である。

【図 11】 本発明の第 5 の実施の形態の説明図である。

【図 12】 本発明の第 5 の実施の形態の機能ブロック図である。

【図 13】 音源位置と i_{max} との関係説明図である。

【図 14】 本発明の第 6 の実施の形態の説明図である。

【図 15】 本発明の第 7 の実施の形態の説明図である。

【図 16】 本発明の第 7 の実施の形態の機能ブロック図である。

【図 17】 本発明の第 8 の実施の形態の説明図である。

【図 18】 本発明の第 9 の実施の形態の説明図である。

【図 19】 本発明の第 10 の実施の形態の説明図である。

【図 20】 従来例のエコーキャンセラの説明図である。

【図 21】 従来例のマイクロホンアレイを用いたエコーキャンセラの説明図である。

【図 22】 従来例の音源位置検出及び目的音強調処理の説明図である。

【符号の説明】

1-1~1-n マイクロホン

2-1~2-n フィルタ

3 加算器

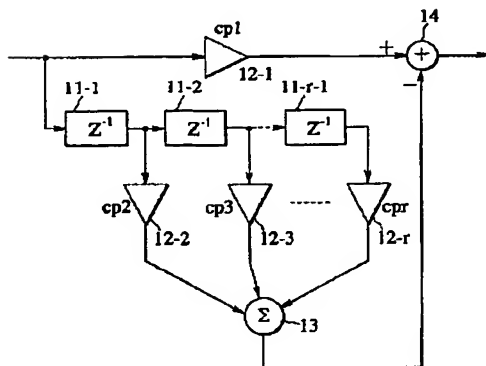
4 フィルタ係数計算部

5 話者（目的音源）

6 スピーカ（雑音源）

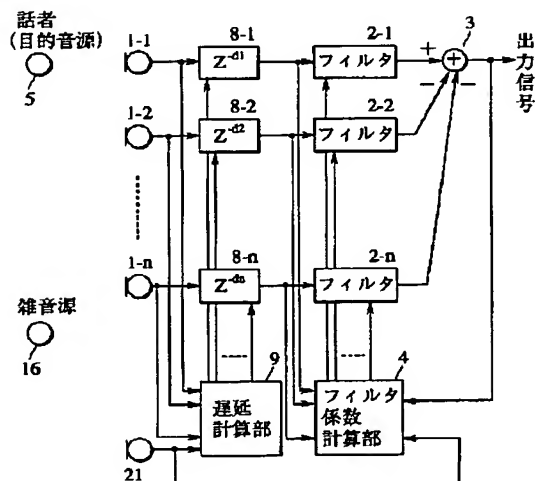
【図 2】

フィルタの説明図



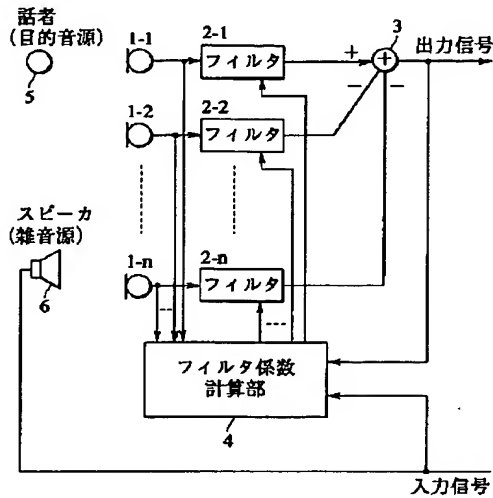
【図 6】

本発明の第4の実施の形態の説明図



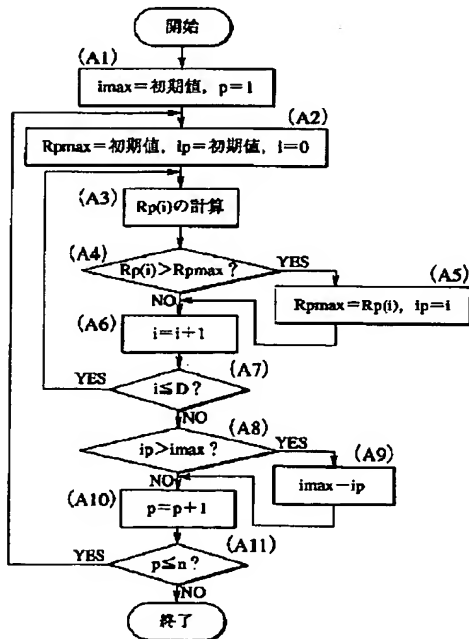
【図 1】

本発明の第1の実施の形態の説明図



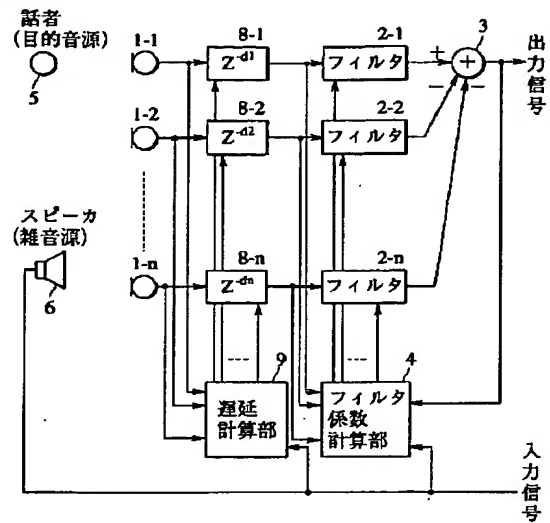
【図 4】

本発明の第2の実施の形態に於ける遅延計算部の処理フローチャート



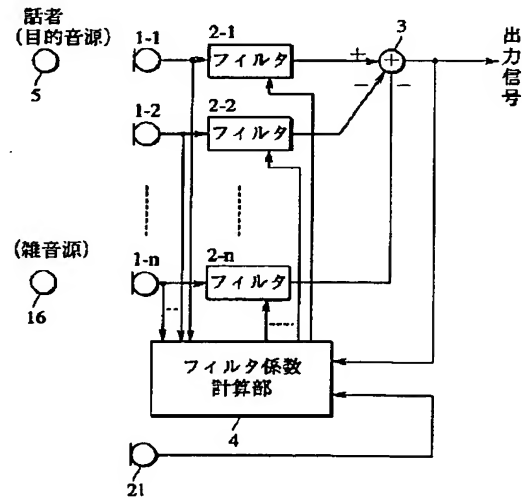
【図 3】

本発明の第2の実施の形態の説明図



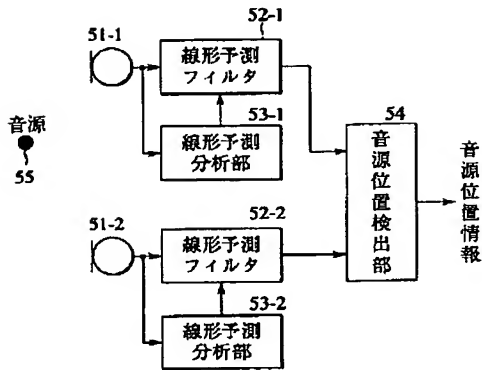
【図 5】

本発明の第3の実施の形態の説明図



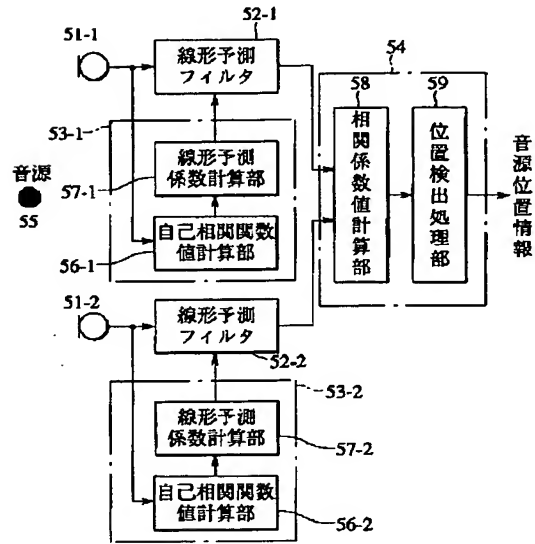
【図 1 1】

本発明の第5の実施の形態の説明図



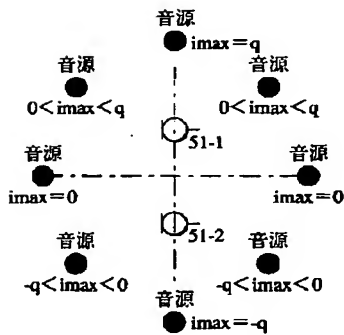
【図 1 2】

本発明の第5の実施の形態の機能ブロック図



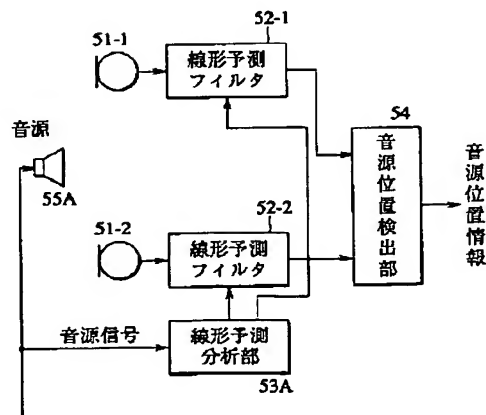
【図 1 3】

音源位置とimaxとの関係説明図



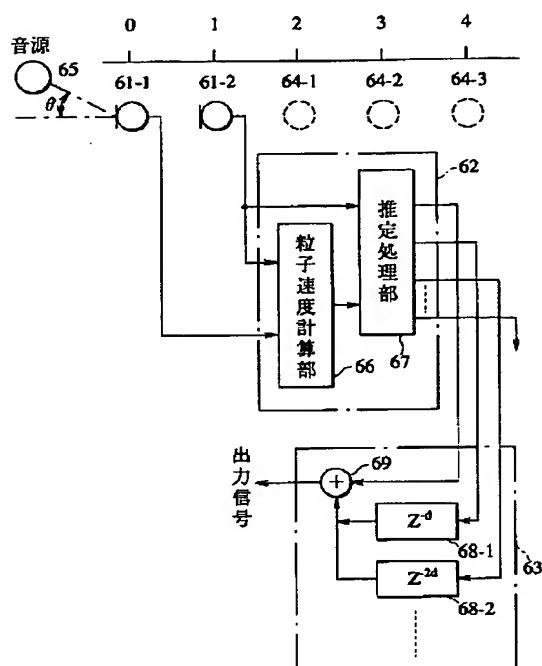
【図 1 4】

本発明の第6の実施の形態の説明図



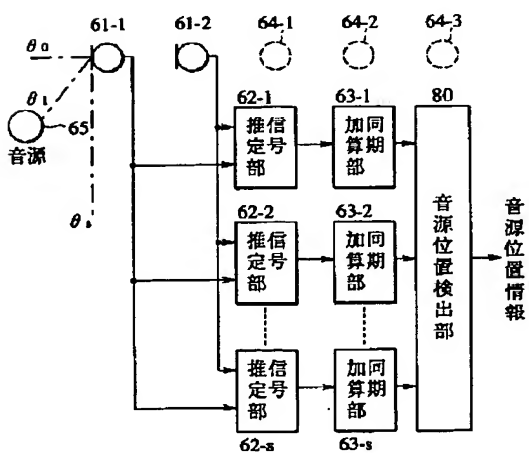
【圖 16】

本発明の第7の実施の形態の機能ブロック図



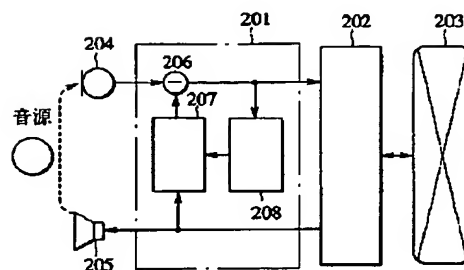
【图 18】

本発明の第9の実施の形態の説明図



【图 20】

従来例のエコーキャンセラの説明図



【圖 2 2】

【図 2 1】

Figure 1 is a block diagram of a speech processing system. On the left, a circle labeled '話者' (Speaker) is connected to a series of input nodes labeled 214-1, 214-2, ..., 214-n. These nodes feed into a central processing block 211. Inside block 211, the inputs pass through a series of stages labeled 217-1, 217-2, ..., 217-n. The output of the final stage 217-n is connected to a node 216. Node 216 is also connected to a block 212 on the right. A feedback path from block 212 goes through a block labeled 'フィルタ係数計算部' (Filter Coefficient Calculation Unit) 218, which then feeds back into the input stages of block 211. The output of block 211 is also connected to a speaker icon labeled 215 at the bottom.

